DNA Sequence, 1999, Vol. 10(1), pp. 37–42
Reprints available directly from the publisher
Photocopying permitted by license only

Short Communication

A Notable Example of an Evolutionary Conserved Gene: Studies on a Putative DNA Helicase TIP49

YUMIKO KUROKAWA[a*], MASATO KANEMAKI[a], YASUTAKA MAKINO[a] and TAKA-AKI TAMURA[b†]

[a]*Department of Biology, Faculty of Science, Chiba University, Yayoi-cho, Inage-ku, Chiba 263–8522, Japan and* [b]*CREST Japan Science and Technology Corporation*

(Received 17 June, 1998)

TIP49a (just called as simply TIP49 in previous reports [Kanemaki *et al.*, 1997; Makino *et al.*, 1998]) was found in a rat nuclear protein complex that included the TATA-binding protein. TIP49a possesses multiple sequence motifs for ATPase and DNA helicase. Since TIP49a structurally resembles prokaryotic DNA helicase RuvB, TIP49a is presumed to be a putative DNA helicase. We demonstrated TIP49a-related gene(s) in variety organisms from human to archaea. Amino acid identities expressed as aligned scores of human, yeast, and *A. fulgidus* TIP49a gene counterparts to the rat sequence were 99, 67, and 46, respectively. Strikingly, two homologous regions of mammalian TIP49a and bacterial RuvB exhibited an aligned score of 17–38. We demonstrated that the eukaryotic TIP49a counterparts were immunologically conserved. These lines of evidence show that the TIP49a gene is a notable example of a highly conserved gene among organisms. An extensive homology search revealed another class of TIP49-related gene in the eukaryotes, designated as TIP49b. Moreover, a phylogenetical study suggested that archaeal TIP49 genes belong to the TIP49b ancestor but not to the TIP49a one and that TIP49a evolved from TIP49b in accordance with divergence of archaea and eukarya. The TIP49 gene family is thought to play a fundamental role in a biological activity.

Keywords: DNA helicase, TIP49, RuvB, molecular evolution, aligned score

TIP49 (TBP-interacting protein 49) was originally isolated as a 456-amino acid (a.a.) component of TBP (TATA-binding protein)-bound nuclear protein complexes purified from rat liver nuclear extracts (Kanemaki *et al.*, 1997). In this study, we changed the nomenclature of this gene to TIP49a because the present work clarified the existence of another TIP49-related gene in eukaryotic genomes (see below). TIP49a has lots of structural features such as Walker A/B (Walker *et al.*, 1982) and some helicase motifs observed in ATPase/GTPase and/or DNA/RNA helicases (Gorbalenya and Koonin, 1993). Moreover, it exhibited a weak similarity with AAA (ATPases associated with various cellular activities) family proteins (Confalonieri and Duguet, 1995). TIP49a has two nuclear localizing signals, one starting at 229 (HKKK) and the other at 265 (KPKK). Actually, a human homolog of TIP49a was localized in the nucleus and appeared as foci (Makino *et al.*, 1998). Thus, TIP49a is catego-

* Y. Kurokawa and M. Kanemaki contributed equally to this study. M.K. is a research fellow of The Japan Society For The Promotion Of Science.
† Corresponding author phone: (+81)43–290–2823, fax: (+81)43–290–2824 E-mail: btamura@nature.s.chiba-u.ac.jp

TABLE I Conservation of TIP49-related proteins in various organisms[1]

	aligned score[2]				
	human	*yeast*[3]	*archaea*[4]	*bacteria*[5]	
TIP49 relatives	TIP49a(100)	YDR190c(67)	TIP49(46)	RuvB[#4](14)	
				region 1 (30)	region 2 (17)
				RuvB[#6](19)	
				region 1 (38)	region 2 (24)
(recombination)	Rad51 (100)	Rad51 (66)	RadA (43)	RecA(15) (19)[7]	
(transcription)	RNA pol II (100)	RNA pol B (53)	RNA pol (39)	RNA pol β' (18)	
	TBP (100)	TBP (60)	TFIID (33)	Sigma factor (<6)	
(replication)	DNA pol α (100)	DNA pol α (29)	N.D.	DNA pol III (8)	

1; Identities of human TIP49a, Rad51, TBP (TATA-binding protein), RNA pol II (the largest subunit of RNA pol II), and DNA pol α (catalytic subunit of DNA polymerase α) with their functional or structural counterparts of various organisms were analyzed. 2; Aligned score was calculated by the CLUSTAL W program and indicated in parentheses. The score value 100 means the same sequence. 3; *S. cerevisiae*, 4; *A. fulgidus*. 5; *E. coli* 6; *T. thermophilus*. 7; Value when only homologous region was searched. #; Two homologous regions (see Fig.1) were also independently calculated. N.D.; Homologous gene has not been assigned.

rized as a nuclear protein. Our previous study demonstrated an unusually high sequence similarity of TIP49a to prokaryotic RuvB, which is denoted as a recombination factor, is involved in branch migration of the Holliday structure (an intermediate of homologous recombination), and possesses an intrinsic DNA helicase activity (Iwasaki *et al.*, 1992; Tsaneva *et al.*, 1993). Based on the available data, we specified TIP49a as a putative DNA helicase. In our most recent publication, we reported the isolation of a human counterpart of rat TIP49a (Makino *et al.*,1998) and noticed that the rat and human genes were 99.8% identical with only one a.a. substitution, implying that the TIP49a equivalent genes are evolutionary conserved.

To obtain additional solid evidence to support the above hypothesis, we analyzed the sequence similarities of TIP49a counterparts in various organisms. We found entire sequences of TIP49a counterparts in *C. elegans*, *S. cerevisiae* (Goffeau et al., 1997) and also in archaea (*A. fulgidus*) (Klenk *et al.*, 1997). Fig. 1 demonstrates the sequence alignment of TIP49a homologs. The a.a. sequence identities expressed as aligned scores of the human, nematoda, yeast, and archaea proteins were 99, 56, 67, and 46 to rat TIP49a, respectively. Surprisingly, rat TIP49a gave significantly high aligned scores to one of the bacterial DNA helicases RuvB (Shinagawa *et al.*, 1988; Tong and Wetmur, 1996) as 14 (*E. coli*) and 19 (*T. thermophilus*). When we focused on two homologous regions (regions 1 and 2, Fig. 1) between rat TIP49a and *T. thermophilus* RuvB, the aligned scores were 38 and 24 for regions 1 and 2, respectively. In general, these high homology scores were unusual. Table I shows sequence identities between mammalian proteins and their yeast, archaeal, and bacterial functional or structural homologs with respect to fundamental factors/enzymes in nucleic acid dynamics that include Rad51 for recombination (Shinohara *et al.*, 1993), RNA polymerase II (RNA pol II) largest subunit (Allison *et al.*, 1985) and TBP for transcription (Horikoshi *et al.*, 1989), and DNA polymerase α (DNA pol α) (Wong *et al.*, 1988) for

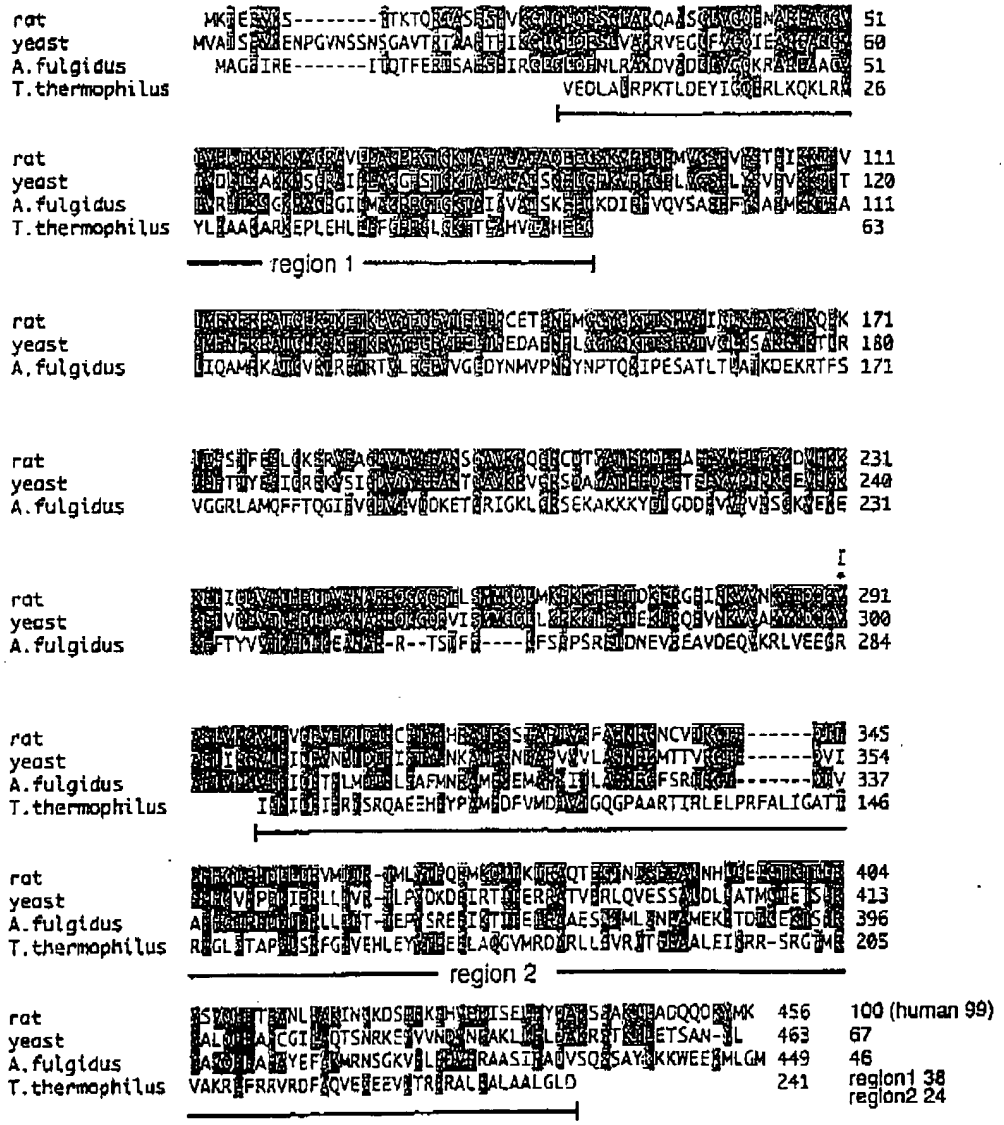

FIGURE 1 Sequence alignment of TIP49a homologs in various organisms Sequences from human TIP49a and archaea TIP49b (TIP49 of *A. fulgidus*) are aligned with rat TIP49a by CLUSTAL W program (Thompson *et al.*, 1994). An asterisk indicates an a.a. substitution (Val to Ile at a.a. 291) in the human counterpart. Sequences of *T. Thermophilus* RuvB were also aligned for two homologous regions: regions 1 and 2. Identical amino acids at least between two sequences are indicated by shadowing. The a.a. sequence identities compared to rat TIP49a are indicated in the bottom line as aligned scores according to the program

replication. As obviously notable, we were not able to observe any similarity between the human DNA pol α and the bacterial one. A counterpart of DNA pol α has not been denoted in the archaea's genome. Though essential mammalian transcription factor TBP showed 33% identity to the archaeal counterpart, it did not exhibit any homology to the sigma factor (a bacterial transcription commitment factor). The aligned score of the RNA pol II to the bacterial RNA polymerase β′ was 18. However, the largest subunit is just one component in the multiple protein complex comprising functional RNA pol II. Rad51 exhibited relatively high score (43) to its archaeal counterpart, whereas the score for mammalian Rad51 to *E. coli* RecA was 15 even though Rad51 has been established to be a functional homolog of RecA. Sequence identities of human TIP49a to archaeal TIP49 was also high (aligned score 46), and those between human TIP49a and bacterial RuvB were at least as high as those for Rad51 as described above and indicated in Table I. From these data, it is evident that TIP49a is a one of the most conserved nuclear dynamics-related proteins.

We carried out immunoblotting of cellular proteins from several eukaryotes by using rat TIP49a polyclonal antibody (Fig. 2). This antibody specifically detected TP49a protein in the rat liver nuclear extract (Fig. 2, lane 2). Interestingly, the antibody also gave signals for proteins from heterologous organisms, whose sizes were nearly equivalent to the rat one. The size of the yeast protein was slightly larger than the rat one (Fig. 2, lane 5). This size difference is well consistent with sequence data because the yeast TIP49a homolog is 1 kDa larger than the rat protein (Fig. 1). As described below, we identified another TIP49-related gene: TIP49b. However, we detected only a single band in the rat extract (Fig. 2). Hence, the immunoblotting bands in Fig. 2 are thought to represent TIP49a homologs and not TIP49b. We suggest that TIP49a counterparts in the eukaryotes are antigenically conserved well.

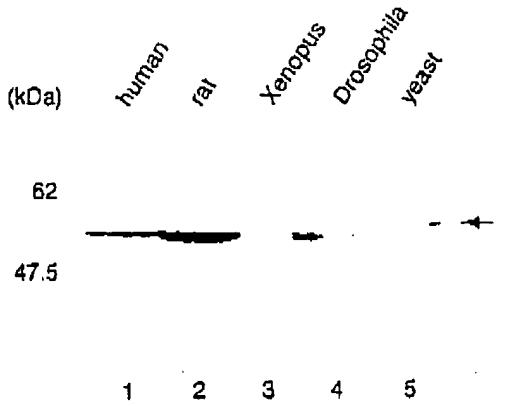

FIGURE 2 Crossreactivity of rat TIP49a antibody to proteins from various eukaryotes Immunoblotting of cell extracts (20 μg) from various organisms was carried out by the alkaline phosphatase method (Promega) employing polyclonal antibody against rat TIP49a. Arrowhead indicates the position of TIP49a proteins. Lane 1, HeLa cell nuclear extract; lane 2, rat liver nuclear extract; lane 3, egg extract of *X. laevis*; lane 4, nuclear extract of *D. melanogaster embryo*; lane 5, extract of *S. cerevisiae*

In the course of a database search for TIP49a-related genes, we found that yeast and nematoda contained another TIP49-related gene in addition to the TIP49a counterpart. We designated this gene as TIP49b. This fact suggested to us the existence of a TIP49 gene family. We drew a phylogenetical tree using available complete TIP49 sequences according to the CLUSTAL W program (Thompson *et al.*, 1994) (Fig. 3). We obtained sequence data of partial but significant length of TIP49-related genes from chicken (TIP49a: Takeda, S., personal communication) and human (TIP49b: Kanemaki, M., unpublished observation), and included them in the phylogenetical tree. The results clearly indicated that TIP49 family genes could be categorized into two branches: TIP49a and -b (Fig. 3). The originally isolated rat TIP49 gene was included in the TIP49a branch. From data for fragmented EST clones, we suppose that the mouse genome also contains a couple of TIP49-related genes. These lines of evidence suggest that eukaryotes com-

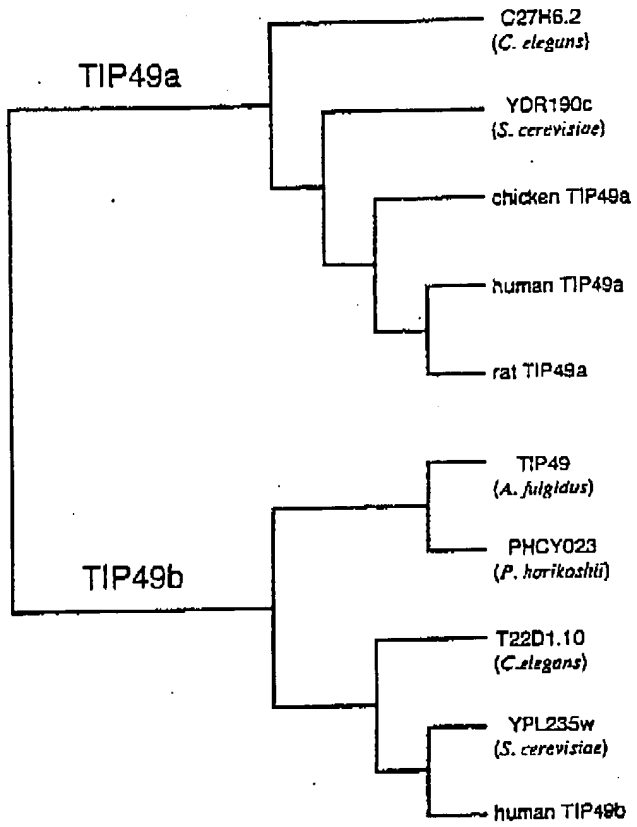

FIGURE 3 Phylogenetical tree of TIP49 family genes Complete a.a. sequences of TIP49a (rat), YDR190c and YPL235W (*S. cerevisiae*), C27H6.2 and T22D1.10 (*C. elegans*), TIP49 (*A. fulgidus*), and PHCY023 (*P. horikoshii*) and partial sequences of chicken TIP49a (Takeda, S., personal communication) and human TIP49b (Karumaki, M., unpublished observation) were analyzed. Calculation was done by the CLUSTAL W program

monly harbor a couple of TIP49-related genes. We found that both TIP49a and -b sequences of *C. elegans* were more distant from the mammalian TIP49 gene family than from the yeast one. We do not know the reason for this fact.

Archaea genome projects revealed that *A. fulgidus* and *P. horikoshii* had only one copy of the TIP49 gene. Phylogenetical analyses indicated that these archaeal TIP49-related genes belong to the TIP49b branch (Fig. 3). Most interestingly, prokaryotic RuvB was calculated to be much closer to the TIP49b branch than to the TIP49a one. If we use the complete human TIP49b sequence for the calculation, sequence identity between RuvB and TIP49 becomes elevated much more. We thus hypothesize that the TIP49b ancestor gene is the prototype of TIP49 gene family, and was transmitted to bacteria and archaea. Then, the TIP49a gene diverged from an archaeal TIP49b ancestor when eukaryotes evolved from archaea.

Determination of higher structure and function for TIP49a/b remain to be done clarify the above issues. Although we consider TIP49a to be

a putative DNA helicase, its biological function is unknown. Since TIP49a is detected ubiquitously, the TIP49a gene must play a fundamental and essential role in eukaryotes and archaea. TIP49a may be involved in transcription because it is included in the TBP-containing complex. Another general transcription factor, TFIIH is known to possess multiple DNA helicase subunits (Svejstrup *et al.*, 1996). Alternatively, TIP49a may be classified as a recombination factor since it is significantly similar to bacterial RuvB in structure.

ABBREVIATIONS USED

TBP: TATA-binding protein, TIP: TBP-interacting protein, a.a.: amino acid; *A. fulgidus, Archaeoglobus. fulgidus; T. thermophilus, Thermusaquaticus thermophilus; P. horikoshii, Pyrococcus horikoshii*

Acknowledgements

We thank Dr. S. Takeda (Kyoto University) for giving us the chicken TIP49a sequence data, Drs. F. Hirose (Aichi Cancer Center Research Institute) and H. Ueda (National Institute of Genetics) for providing *Drosophila* extracts, and Drs. T. Ogawa (National Institute of Genetics) and H. Shinagawa (Osaka University) for valuable discussions. A part of this work was supported by a grant from *The Japanese Ministry of Education, Science, Sports, and Culture.*

References

Allison, L. A., Moyle, M., Shales, M. and Ingles, C. J. (1985) "Extensive homology among the largest subunits of eukaryotic and prokaryotic RNA polymerases", *Cell*, 42, 599–610.

Confalonieri, F. and Duguet, M. (1995) "A 200-amino acid ATPase module in search of a basic function", *Bioessays*, 17, 639–650.

Goffeau, A., Aert, R. *et al.* (1997) "The yeast genome directory *Nature*, 387 (*suppl.*), 1–105.

Gorbalenya, A. E. and Koonin, E. V. (1993) "Helicases: amino acid sequence comparisons and structure-function relationships", *Curr. Opin. Struct. Biol.*, 3, 419–429.

Horikoshi, M., Wang, C. K., Fujii, H., Cromlish, J. A., Weil, P. A. and Roeder, R. G. (1989) "Cloning and structure of a yeast gene encoding a general transcription initiation factor TFIID that binds to the TATA box", *Nature*, 341, 299–303.

Iwasaki, H., Takahagi, M., Nakata, A. and Shinagawa, H. (1992) "*Escherichia coli* RuvA and RuvB proteins specifically interact with Holliday junctions and promote branch migration", *Genes Dev.*, 6, 2214–2220.

Kanemaki, M., Makino, Y., Yoshida, T., Kishimoto, T., Koga, A., Yamamoto, K., Yamamoto, M., Moncollin, V., Egly, J.-M., Muramatsu, M. and Tamura, T. (1997) "Molecular Cloning of a Rat 49-kDa TBP-interacting protein (TIP49) that is highly homologous to the bacterial RuvB", *Biochem. Biophys. Res. Commun.*, 235, 64–68.

Klenk, H.-P., Clayton, R. A., *et al.* (1997) "The complete genome sequence of the hyperthermophilic, sulphate-reducing archaeon *Archaeoglobus fulgidus*", *Nature*, 390, 364–370.

Makino, Y., Mimori, T., Koike, C., Kanemaki, M., Kurokawa, Y., Inoue, S., Kishimoto, T. and Tamura, T. (1998) "TIP49, homologous to the bacterial DNA helicase RuvB, acts as an autoantigen in human", *Biochem. Biophys. Res. Commun.*, 245, 819–823.

Shinagawa, H., Makino, K., Amemura, M., Kimura, S., Iwasaki, H. and Nakata, A. (1988) "Structure and regulation of the *Escherichia coli ruv* operon involved in DNA repair and recombination", *J. Bacteriol.*, 170, 4322–4329.

Shinohara, A., Ogawa, H., Matsuda, Y., Ushio, N., Ikeo, K. and Ogawa, T. (1993) "Cloning of human, mouse and fission yeast recombination genes homologous to RAD51 and recA", *Nature genetics*, 4, 239–243.

Svejstrup, J. Q., Vichi, P. and Egly, J.-M. (1996) "The multiple roles of transcription/repair factor TFIIH", *Trends Biochem. Sci.*, 21, 346–350.

Thompson, J. D., Higgins, D. G. and Gibson, T. J. (1994) "CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice", *Nucl. Acids Res.*, 22, 4673–4680.

Tong, J. and Wetmur, J. G. (1996) "Cloning, sequencing, and expression of ruvB and characterization of RuvB proteins from two distantly related thermophilic eubacteria", *J. Bacteriol.*, 178, 2695–2700.

Tsaneva, I. R., Muller, B. and West, S. C. (1993) "RuvA and RuvB proteins of *Escherichia coli* exhibit DNA helicase activity *in vitro*", *Proc. Natl. Acad. Sci. USA*, 90, 1315–1319.

Walker, J. E., Saraste, M., Runswick, M. J. and Gay, N. J. (1982) "Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotides binding fold", *EMBO J.*, 1, 945–951.

Wong, S. W., Wahl, A. F., Yuan, P.-M., Arai, N., Pearson, B. E., Arai, K., Korn, D., Hunkapiller, M. W. and Wang, T.S.-F. (1988) "Human DNA polymerase α gene expression is cell proliferation dependent and its primary structure is similar to both prokaryotic replicative DNA polymerases", *EMBO J.*, 7, 37–47.